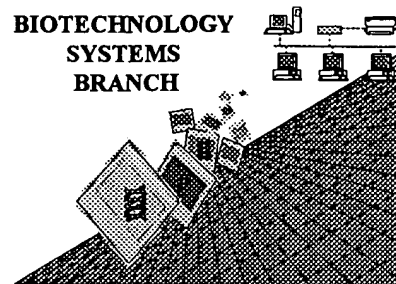# RAW SEQUENCE LISTING
ERROR REPORT

BIOTECHNOLOGY
SYSTEMS
BRANCH

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number: _08/846,606B_

Art Unit / Team No. : _1636_

Date Processed by STIC: _8/31/98_

THE ATTACHED PRINTOUT EXPLAINS THE **ERRORS DETECTED.**

PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,

2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY

THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.

IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:

ARTI SHAH 703-308-4212

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/846,606B*

DATE: 09/01/98
TIME: 15:46:12

INPUT SET: S28371.raw

This Raw Listing contains the General Information Section and those Sequences containing ERRORS.

Does Not Comply
Corrected Diskette Needed

```
1                                    SEQUENCE LISTING
2
3    (1)     General Information:
4            (i) APPLICANT:  Schering Corporation
5            (ii) TITLE OF INVENTION: Thioredoxin/Heterologous Protein Bacterial Expression
6            (iii) NUMBER OF SEQUENCES: 14
7            (iv) CORRESPONDENCE ADDRESS:
8    (A) ADDRESSEE: Schering-Plough Corporation
9    (B) STREET: 2000 Galloping Hill Road
10   (C) CITY: Kenilworth
11   (D) STATE: New Jersey
12   (E) COUNTRY: U.S.A.
13   (F) ZIP : 07033-0530
14          (v) COMPUTER READABLE FORM:
15   (A) MEDIUM TYPE: diskette
16   (B) COMPUTER: Apple Macintosh
17   (C) OPERATING SYSTEM: Macintosh 7.5.3
18   (D) SOFTWARE: Microsoft Word 5.1a
19      (vi) CURRENT APPLICATION DATA:
20   (A) APPLICATION NUMBER:
21   (B) FILING DATE:
22      (vii) PRIOR APPLICATION DATA
23   (A) APPLICATION NUMBER: US 60/011,606
24   (B) FILING DATE: 30-APR-1996
25     (viii) ATTORNEY/AGENT INFORMATION:
26   (A) NAME: Thampoe, Immac J..
27   (B) REGISTRATION NUMBER: 36,322
28   (C) REFERENCE DOCKET NUMBER: JB0600Q
29
```

ERRORED SEQUENCES FOLLOW:

```
41   (2) INFORMATION FOR SEQ ID NO:2:
42          (i) SEQUENCE CHARACTERISTICS:
-->  43   (A) LENGTH: 47 bases
44   (B) TYPE: nucleic acid
45   (C) STRANDEDNESS: single
46   (D) TOPOLOGY: linear
47          (ii) MOLECULE TYPE: DNA
48          (xi) SEQUENCE DESCRIPTION:   SEQ ID NO:2:
49
50   GCACCCAACA TGCAAGGATC CTTACGCCAG ATTAGCATCG AGGAACT
```

Per 1.822 of sequence Rules insert the cumulative base Total at right margin

→ 47 margin

of level
line

```
       51

       52   (2)  INFORMATION FOR SEQ ID NO:3:
       53        (i)  SEQUENCE CHARACTERISTICS:
-->    54   (A) LENGTH: 336 base pairs
       55   (B) TYPE: nucleic acid
       56   (C) STRANDEDNESS: single
       57   (D) TOPOLOGY: linear
       58        (ii) MOLECULE TYPE: DNA
       59        (xi) SEQUENCE DESCRIPTION:   SEQ ID NO:3:
       60
       61   ATGAGCGATA AAATTATTCA CCTGACTGAC GACAGTTTTG ACACGGATGT 50
       62   ACTCAAAGCG GACGGGGCGA TCCTCGTCGA TTTCTGGGCA GAGTGGTGCG 100
       63   GTCCGTGCAA AATGATCGCC CCGATTCTGG ATGAAATCGC TGACGAATAT 150
       64   CAGGGCAAAC TGACCGTTGC AAAAACTGAAC ATCGATCAAA ACCCTGGCAC 200
       65   TGCGCCGAAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA 250
       66   ACGGTGAAGT GGCGGCAACC AAAGTGGGTG CACTGTCTAA AGGTCAGTTG 300
       67   AAAGAGTTCC TCGATGCTAA TCTGGCGTAA GGATCC
       68

       69   (2)  INFORMATION FOR SEQ ID NO:4:
       70        (i)  SEQUENCE CHARACTERISTICS:
-->    71   (A) LENGTH: 336 bases pairs
       72   (B) TYPE: nucleic acid
       73   (C) STRANDEDNESS: single
       74   (D) TOPOLOGY: linear
       75        (ii) MOLECULE TYPE: cDNA
       76        (xi) SEQUENCE DESCRIPTION:   SEQ ID NO:4:
       77
       78   ATGAGCGATA AAATTATTCA CCTGACTGAC GACAGTTTTG ACACGGATGT 50
       79   ACTCAAAGCG GACGGGGCGA TCCTCGTCGA TTTCTGGGCA GAGTGGTGCG 100
       80   GTCCGTGCAA AATGATCGCC CCGATTCTGG ATGAAATCGC TGACGAATAT 150
       81   CAGGGCAAAC TGACCGTTGC AAAAACTGAAC ATCGATCAAA ACCCTGGCAC 200
       82   TGCGCCGAAA TATGGCATCC GTGGTATCCC GACTCTGCTG CTGTTCAAAA 250
       83   ACGGTGAAGT GGCGGCAACC AAAGTGGGTG CACTGTCTAA AGGTCAGTTG 300
       84   AAAGAGTTCC TCGAGGCTAA TCTGGCGTAA GGATCC
       85

       86   (2)  INFORMATION FOR SEQ ID NO:5:
       87        (i)  SEQUENCE CHARACTERISTICS:
-->    88   (A) LENGTH: 81 bases pairs
       89   (B) TYPE: nucleic acid
       90   (C) STRANDEDNESS: single
       91   (D) TOPOLOGY: linear
       92        (ii) MOLECULE TYPE: cDNA
       93        (xi) SEQUENCE DESCRIPTION:   SEQ ID NO:5:
       94
       95   GAT AAT ATT CTG GCT GGT TCT GGT TCT GGT GAT GAC GAT GAC AAG45
       96   Asp Asn Asn Leu Ala Gly Ser Gly Ser Gly Asp Asp Asp Asp Lys
       97   1             5                   10                  15
       98
       99
      100   GGT CCT GTT CCG CCG TCT ACC GCT CTG CGT GAG CTC
```

INPUT SET: S28371.raw

```
101   Gly Pro Val Pro Pro Ser Thr Ala Leu Arg Glu Leu
102                   20                  25
103
104
```

```
121   (2) INFORMATION FOR SEQ ID NO:7:
122         (i) SEQUENCE CHARACTERISTICS:
123   (A) LENGTH: 52 bases
124   (B) TYPE: nucleic acid
125   (C) STRANDEDNESS: single
126   (D) TOPOLOGY: linear
127       (ii) MOLECULE TYPE: DNA
128       (xi) SEQUENCE DESCRIPTION:  SEQ ID NO:7:
129
130   GAAGGAGGCT GATTAAATGG GTCCGGTTCC GCCGTCTACC GCTCTGGAGC 50
131   TC
132
```

```
133   (2) INFORMATION FOR SEQ ID NO:8:
134         (i) SEQUENCE CHARACTERISTICS:
135   (A) LENGTH: 18 bases
136   (B) TYPE: nucleic acid
137   (C) STRANDEDNESS: single
138   (D) TOPOLOGY: linear
139       (ii) MOLECULE TYPE: DNA
140       (xi) SEQUENCE DESCRIPTION:  SEQ ID NO:8:
141
142       AAGGAGGCTG ATTAAATG
143
```

```
144   (2) INFORMATION FOR SEQ ID NO:9:
145         (i) SEQUENCE CHARACTERISTICS:
146   (A) LENGTH: 17 bases
147   (B) TYPE: nucleic acid
148   (C) STRANDEDNESS: single
149   (D) TOPOLOGY: linear
150       (ii) MOLECULE TYPE: DNA
151       (xi) SEQUENCE DESCRIPTION:  SEQ ID NO:9:
152
153       AAGGAGGCTG ATTAATG
154
```

```
155   (2) INFORMATION FOR SEQ ID NO:10:
156         (i) SEQUENCE CHARACTERISTICS:
157   (A) LENGTH: 14 bases
158   (B) TYPE: nucleic acid
159   (C) STRANDEDNESS: single
160   (D) TOPOLOGY: linear
161       (ii) MOLECULE TYPE: DNA
162       (xi) SEQUENCE DESCRIPTION:10:
163
164   AAGGAGGTTT AATG
165
```

add: "SEQ ID NO" (this error appears in subsequent sequences) ← insert total

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/08/846,606B*

DATE: 09/01/98
TIME: 15:46:15

INPUT SET: S28371.raw

```
166       (2) INFORMATION FOR SEQ ID NO:11:
167            (i) SEQUENCE CHARACTERISTICS:
168       (A) LENGTH: 6 amino acids
169       (B) TYPE: amino acid
170       (D) TOPOLOGY: linear
171            (ii) MOLECULE TYPE: peptide
-->  172       (xi) SEQUENCE DESCRIPTION:11)   see Seq 10
173
174       Leu    Asp    Ala    Asn    Leu    Ala
175       1                           5
```

```
176       (2) INFORMATION FOR SEQ ID NO:12:
177            (i) SEQUENCE CHARACTERISTICS:
-->  178       (A) LENGTH: 21 bases
179       (B) TYPE: nucleic acid
180       (C) STRANDEDNESS: single
181       (D) TOPOLOGY: linear
182            (ii) MOLECULE TYPE: DNA
183            (xi) SEQUENCE DESCRIPTION:12:
184
185       CTC    GAT    GCT    AAT    CTG    GCG    TAA         ↙ total
186       Leu    Asp    Ala    Asn    Leu    Ala
187       1                           5
188
189
```

```
190       (2) INFORMATION FOR SEQ ID NO:13:
191            (i) SEQUENCE CHARACTERISTICS:
-->  192       (A) LENGTH: 21 bases
193       (B) TYPE: nucleic acid
194       (C) STRANDEDNESS: single
195       (D) TOPOLOGY: linear
196            (ii) MOLECULE TYPE: DNA
197            (xi) SEQUENCE DESCRIPTION:13)
198
199       CTC    GAG    GCT    AAT    CTG    GCG    TAA         ↙ total
200       Leu    Glu    Ala    Asn    Leu    Ala
201       1                           5
202
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/846,606B*

INPUT SET: S28371.raw

Line	Error	Original Text
43	Entered (47) and Calc. Seq. Length (0) differ	(A) LENGTH: 47 bases
54	Entered (336) and Calc. Seq. Length (300) differ	(A) LENGTH: 336 base pairs
71	Entered (336) and Calc. Seq. Length (300) differ	(A) LENGTH: 336 bases pairs
88	Entered (81) and Calc. Seq. Length (45) differ	(A) LENGTH: 81 bases pairs
123	Entered (52) and Calc. Seq. Length (50) differ	(A) LENGTH: 52 bases
135	Entered (18) and Calc. Seq. Length (0) differ	(A) LENGTH: 18 bases
146	Entered (17) and Calc. Seq. Length (0) differ	(A) LENGTH: 17 bases
157	Entered (14) and Calc. Seq. Length (0) differ	(A) LENGTH: 14 bases
172	Mandatory Value Not Present	(xi) SEQUENCE DESCRIPTION SEQ ID NO:11:
178	Entered (21) and Calc. Seq. Length (0) differ	(A) LENGTH: 21 bases
192	Entered (21) and Calc. Seq. Length (0) differ	(A) LENGTH: 21 bases